

Evidence of statistical epistasis between DISC1, CIT and NDEL1 impacting risk for schizophrenia: biological validation with functional neuroimaging

Kristin K. Nicodemus · Joseph H. Callicott · Rachel G. Higier · Augustin Luna ·
Devon C. Nixon · Barbara K. Lipska · Radhakrishna Vakkalanka · Ina Giegling ·
Dan Rujescu · David St. Clair · Pierandrea Muglia · Yin Yao Shugart · Daniel R. Weinberger

Received: 29 September 2009 / Accepted: 24 December 2009 / Published online: 19 January 2010
© US Government 2010

Abstract The etiology of schizophrenia likely involves genetic interactions. DISC1, a promising candidate susceptibility gene, encodes a protein which interacts with many other proteins, including CIT, NDEL1, NDE1, FEZ1 and PAFAH1B1, some of which also have been associated with psychosis. We tested for epistasis between these genes in a schizophrenia case–control study using machine learning algorithms (MLAs: random forest, generalized boosted regression and Monte Carlo logic regression). Convergence of MLAs revealed a subset of seven SNPs that were subjected to 2-SNP interaction modeling using likelihood ratio tests for

nested unconditional logistic regression models. Of the ${}^7C_2 = 21$ interactions, four were significant at the $\alpha = 0.05$ level: DISC1 rs1411771–CIT rs10744743 OR = 3.07 (1.37, 6.98) $p = 0.007$; CIT rs3847960–CIT rs203332 OR = 2.90 (1.45, 5.79) $p = 0.003$; CIT rs3847960–CIT rs440299 OR = 2.16 (1.04, 4.46) $p = 0.038$; one survived Bonferroni correction (NDEL1 rs4791707–CIT rs10744743 OR = 4.44 (2.22, 8.88) $p = 0.00013$). Three of four interactions were validated via functional magnetic resonance imaging (fMRI) in an independent sample of healthy controls; risk associated alleles at both SNPs predicted prefrontal cortical inefficiency during the N -back task, a schizophrenia-linked intermediate biological phenotype: rs3847960–rs440299; rs1411771–rs10744743, rs4791707–rs10744743 (SPM5 $p < 0.05$, corrected), although we were unable to statistically replicate the

Electronic supplementary material The online version of this article (doi:10.1007/s00439-009-0782-y) contains supplementary material, which is available to authorized users.

K. K. Nicodemus · J. H. Callicott · R. G. Higier · A. Luna ·
D. C. Nixon · B. K. Lipska · R. Vakkalanka ·
D. R. Weinberger (✉)
Genes, Cognition and Psychosis Program,
Intramural Research Program,
National Institute of Mental Health,
National Institutes of Health,
Room 4S-235, 10 Center Drive,
Bethesda, MD 20892, USA
e-mail: weinberd@mail.nih.gov

K. K. Nicodemus
Wellcome Trust Centre for Human Genetics,
University of Oxford, Roosevelt Road,
Oxford OX3 7BN, UK
e-mail: kristin.nicodemus@well.ox.ac.uk

K. K. Nicodemus
Department of Clinical Pharmacology,
Old Road Campus Research Building,
University of Oxford, Old Road Campus,
Off Roosevelt Drive, Headington, Oxford OX3 7DQ, UK

I. Giegling · D. Rujescu
Section of Molecular and Clinical Neurobiology,
Department of Psychiatry, Ludwig Maximilians
University, Nussbaumstrasse 7, 80336 Munich, Germany

D. St. Clair
Institute of Medical Sciences, University of Aberdeen,
Foresterhill, Aberdeen AB25 2ZD, UK

P. Muglia
GlaxoSmithKline, Via Fleming 4, Verona 37135, Italy

Y. Y. Shugart
Department of Epidemiology, Johns Hopkins Bloomberg School
of Public Health, 615 N Wolfe St, Baltimore, MD 21205, USA

Y. Y. Shugart
Genomic Research Branch, Division of Neuroscience Center,
National Institute of Mental Health, National Institutes
of Health, Room 7191, 6001 Executive Blvd, Bethesda,
MD 20892, USA

interactions in other clinical samples. Interestingly, the CIT SNPs are proximal to exons that encode the DISC1 interaction domain. In addition, the 3' UTR DISC1 rs1411771 is predicted to be an exonic splicing enhancer and the NDEL1 SNP is ~3,000 bp from the exon encoding the region of *NDEL1* that interacts with the DISC1 protein, giving a plausible biological basis for epistasis signals validated by fMRI.

Introduction

A translocation that disrupts the Disrupted in Schizophrenia 1 (*DISC1*) gene was found to segregate with major psychiatric disorders in a large Scottish family (St. Clair et al. 1990), making it an attractive candidate gene for schizophrenia. Two ways in which the translocation may be disease related have been suggested: through disruption of protein-binding domains of the gene product (Millar et al. 2000, 2001) or via haploinsufficiency (Millar et al. 2005). We previously reported an association between SNPs in *DISC1* and schizophrenia and hippocampal structure in a European American sample (Callicott et al. 2005), and other researchers have also reported association between *DISC1* polymorphisms and schizophrenia (Burdick et al. 2008; Cannon et al. 2005; Hennah et al. 2003, 2008; Hodgkinson et al. 2004; Kockelkorn et al. 2004; Liu et al. 2006; Saetre et al. 2008; Schumacher et al. 2009; Song et al. 2008; Thomson et al. 2005; Wood et al. 2007; Zhang et al. 2006, 2007), although not all studies have detected an association (Devon et al. 2001; Sanders et al. 2008; Zhang et al. 2005). Mounting evidence has shown that the gene products of *DISC1*, *PAFAH1B1* (previously known as *LIS1*), *FEZ1*, *NDE1*, *NDEL1* and *CIT* interact directly or indirectly and they may control various aspects of neurodevelopment, lateral ventricle size, behavior, gene expression and function in human and rodent brain (Brandon et al. 2004; Burdick et al. 2008; Clapcote et al. 2007; Feng et al. 2000; Hikida et al. 2007; Kamiya et al. 2006; Li et al. 2007; Lipska et al. 2006; Miyoshi et al. 2003; Morris et al. 2003; Ozeki et al. 2003; Pletnikov et al. 2008; Shu et al. 2004; Tarricone et al. 2004; Taya et al. 2007). Because risk for schizophrenia is thought to be determined by a network of gene–gene and gene–environment interactions, a natural step in determining whether variation within the putative *DISC1* protein pathway influences risk for schizophrenia is to look for evidence for statistical interaction impacting disease risk, followed by biologic validation using an independent biological system, such as functional magnetic resonance imaging (fMRI) of a working memory task, a paradigm that elicits a heritable (Blokland et al. 2008) pattern of prefrontal cortical activity that is related to schizophrenia and to increased genetic risk for schizophrenia (Callicott et al. 2003b). Of the putative *DISC1* interaction

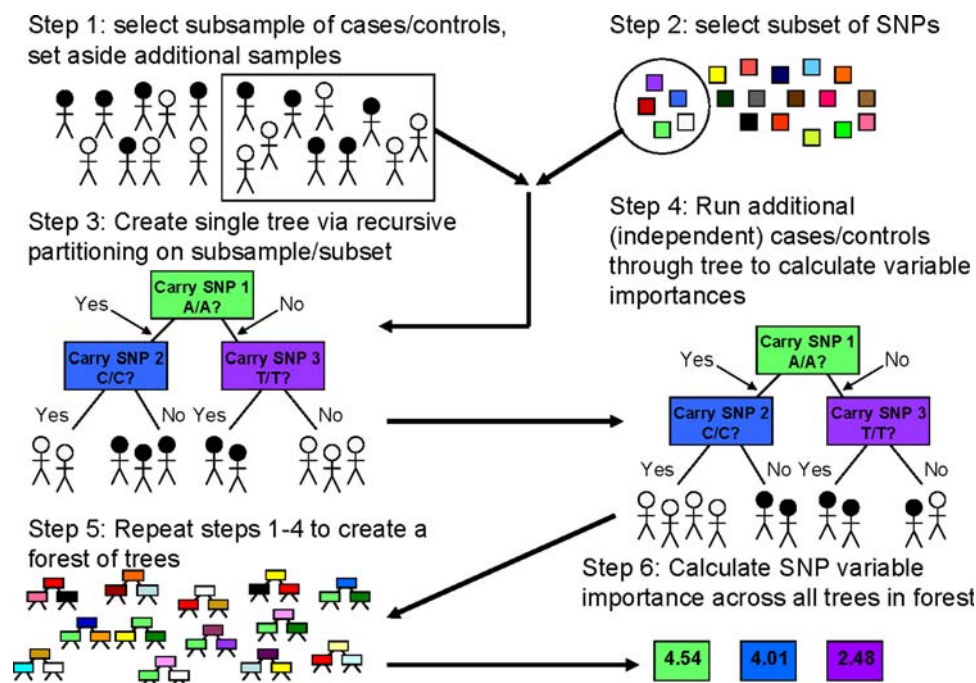
partners, *NDE1* and *NDEL1* have been shown to interact with *DISC1* to increase risk for schizophrenia (Burdick et al. 2008) and a SNP and haplotype in *NDEL1* has been associated with schizophrenia (Tomppa et al. 2009); a SNP within *NDE1* was also associated with schizophrenia in women (Hennah et al. 2007), and one study each has reported no evidence for association between schizophrenia and *NDE1* (Numata et al. 2008)/*NDEL1* (Kähler et al. 2008). Although one study reported association between *FEZ1* and schizophrenia (Yamada et al. 2004), three additional studies did not replicate the association (Hodgkinson et al. 2007; Koga et al. 2007; Tomppa et al. 2009). *CIT* has been associated with bipolar disorder, a disorder thought to share genetic etiology with schizophrenia including association with *DISC1* (Lyons-Warren et al. 2005), although no studies thus far have examined association between *CIT* and schizophrenia.

DISC1 has been considered a developmental hub protein, with many other potential interactors in addition to those noted above (Carmargo et al. 2007). In an effort to explore the potential validity of epistatic interactions between *DISC1* and several of its potential partners, we selected a limited number of variants (50 SNPs) in *DISC1* and five major partners, recognizing that there are likely many other networks that could be selected. We used a case ($N = 289$)–control ($N = 359$) study, applying three methods designed for use with high-dimensional data: random forest (RF; Breiman et al. 1984; Breiman 2001) (Fig. 1), Monte Carlo logic regression (MCLR; Kooperberg et al. 2001; Kooperberg and Ruczinski, 2005) and generalized boosted regression (GBM; Friedman 2001) and sought to biologically validate interactions detected statistically using fMRI in an independent sample of healthy controls.

Materials and methods

Study population and phenotyping

Subjects with schizophrenia ($n = 296$, 230 male, 66 female) were ascertained through current or former patients of NIMH, recruitment nationwide via the National Alliance for the Mentally Ill (NAMI), public advertising, Chestnut Lodge Hospital, St. Elizabeth's Hospital, and physician referrals. Eligibility criteria for participation by cases included: diagnosis of schizophrenia ($N = 245$), schizoaffective disorder ($N = 37$), psychosis not otherwise specified (NOS) ($N = 9$), or schizophreniform disorder ($N = 5$), age 18–65, measured IQ greater than 70, no history of brain damage or neurological disease, no history of significant alcohol or substance abuse, and the ability to give informed consent. For controls ($n = 365$, 174 male, 191 female), the eligibility criteria included: no diagnosis of psychiatric disorder, no family history of psychiatric disorder in a first

Fig. 1 Diagram showing random forest algorithm

degree relative, age 18–65, measured IQ greater than 70, no history of brain damage or neurological disease, and no history of significant alcohol or substance abuse. Healthy subjects were recruited through the NIH Normal Volunteer Office. All subjects gave written, informed consent to participate in the NIMH Genetic Study of Schizophrenia (ClinicalTrials.gov Identifier: NCT00001486).

Methods used in phenotyping for the subjects included in this study have been previously described (Egan et al. 2000). Briefly, all individuals participating in the study were given a structured clinical diagnostic interview by at least one staff psychiatrist. The structured diagnostic interview given to participants is based on the Structured Clinical Interview for DSM-IV for Axis I Disorders Research Version (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II). A second psychiatrist independently reviewed the SCID interview information and psychiatric records. Any disagreements in diagnosis were referred to a third psychiatrist for final diagnosis. Interviewer reliability was assessed with a small number of cases (14–17) within the study by conducting independent interviews. In all cases reviewers agreed on their final diagnosis (Egan et al. 2000).

Sample collection and genotyping

Blood was collected and DNA was extracted using standard methods. Genotyping of all SNPs was performed using the Taqman 5'-exonuclease allelic discrimination assay. Two raters independently manually assessed genotype calls, and discordant genotypes were resolved by

mutual consensus or by assigning the genotype as missing. Genotypes were regularly checked by re-genotyping and have shown 99% reproducibility; in addition, spot accuracy checks have been performed by checking Taqman-generated genotypes against results from double stranded sequencing and have shown >99% agreement.

SNP selection

The selection of 12 SNPs within *DISC1* has been described elsewhere (Callicott et al. 2005) and was focused on coding SNPs and physical distance with a low r^2 threshold to reduce redundant information. SNPs from the *DISC1* interaction partner genes were selected using tagSNP data from the HapMap (<http://www.hapmap.org>) (de Bakker et al. 2005). SNPs were selected based on the minimal number of possible SNPs with an r^2 cutoff value of 0.8 and minor allele frequency >5%, and extended 10 kb 5' and 5 kb 3' of the gene. Numbers of SNPs genotyped per gene were: FEZ1 (13 SNPs), NDEL1 (1 SNP), NDE1 (3 SNPs), CITRON (19 SNPs), PAFAH1B1 (2 SNPs), for a total of 50 SNPs in six genes. Minor allele frequencies, and linkage disequilibrium metrics (D' and r^2) are shown in Supplementary Tables S1–S4.

Data integrity and quality control

Testing of Hardy–Weinberg equilibrium (HWE) with Fisher's exact test was completed for cases and controls separately, using the module GENHW implemented in STATA 8.2 (College Park, TX). Data consistency and error checking were completed before statistical analysis was

conducted, and resulted in two SNPs in CITRON being dropped from further analyses because genotyping call rates were less than 90% (rs904654 and rs203364). Two SNPs in CITRON were found to be mildly out of HWE proportions in controls only (rs278126, p value = 0.043 and rs202983, p value = 0.027); one SNP in PAFAH1B1 and one SNP in NDE1 were also out of HWE in controls (rs7212450; p value = 0.001; rs4781680, p value = 0.012). SNPs out of HWE were retained in the analyses after being examined by laboratory staff; the numbers of SNPs out of HWE in cases and in controls is less than the number expected by chance with $\alpha = 0.05$ for 50 markers.

Statistical methodologies

We used three machine learning algorithms (MLAs) to assess complex interactions between SNPs in six genes with the goal of finding a consensus between the variable importance measures; essentially, a meta-machine learning approach, which may improve prediction versus single algorithm results. When effect sizes are moderate, the resulting empirical variable importance measures from each algorithm may be slightly unstable (Nicodemus et al. 2007; Nicodemus and Malley 2009); therefore, we repeated each analysis 500 times and used the median of these 500 runs to obtain stable estimates of variable importance. We considered a SNP for follow-up if the median variable importance was the top 10 of the distribution of all variable importance measures and had an empirical p value of <0.05 for >1 algorithm. Missing genotypes were inferred using randomForest (Breiman 2001). The maximum number of genotypes inferred for any SNP was 7.6% and the minimum was 1.2%. In addition, all analyses were performed on 500 null replicates where case status had been randomly permuted to evaluate significance of the median observed variable importance measures. To obtain estimates of the effect sizes for two-way interaction and to define interacting SNPs, we performed all-possible two-way interaction modeling via logistic regression among the SNPs considered influential across >1 machine learning algorithm. Significance of the logistic regression models was obtained using a likelihood ratio test (LRT) comparing nested models; a reduced model containing terms modeling the main effects of each SNP and a full model containing the main effects plus interaction terms.

Random forest algorithm

The random forest (RF) algorithm (Breiman et al. 1984; Breiman 2001) as implemented in the R package randomForest relies on binary classification trees as the base or weak learner. To begin, a subset of predictors is randomly sampled; we used the tuneRF function in randomForest to estimate the optimal number of predictors for our

data (N predictors = 7). In addition, a subsample of the observations is selected for tree building; we used subsampling of 63.2% of the total observations. A single tree is created using recursive partitioning of the subsampled predictors on the subsampled observations, where the splitting rule is based on which variable gives the largest decrease in impurity as measured by the Gini Index. The algorithm terminates when no additional variables produce decreases in impurity or when the terminal node size is less than five observations; the process is repeated to create a forest of classification trees; we used a forest size of 5,000 trees. Differences between the observed importance measure calculated on the independent set of the observations (the 36.8% not used to grow the tree) and importance measures obtained after permutation of the genotype labels, averaged across all the trees in the forest containing that predictor, provide an empirical measure of variable importance for each SNP as an interactor in the RF algorithm.

Monte Carlo logic regression

Monte Carlo logic regression (Kooperberg et al. 2001; Kooperberg and Ruczinski 2005), as implemented in the R package LogicReg, is a regression-based method that constructs logic trees instead of classification trees as the base learner. Logic trees are comprised of Boolean combinations of binary independent variables. Allowable Boolean operators in constructing logic trees include AND, OR and NOT. The algorithm creates a forest of logic trees via moves between adjacent trees during modeling which is conducted with a reversible jump Markov chain Monte Carlo (rjMCMC) algorithm. The variable importance measure is a count of how many times a variable is selected to be in a tree across all moves in the chain. Because MCLR accepts only binary predictors, we recoded each SNP into two variables: one with the minor allele coded as dominant and one with the minor allele coded as recessive. We performed MCLR using a burn-in interval of 10,000 and a Markov chain length of 1,000,000. For consistency with RF, the maximum model size for MCLR was set at 2 trees with 4 leaves (8 predictors).

Generalized boosted regression

Boosting, as implemented in the R package gbm, is a stagewise additive expansion of small classification trees to reduce the loss function, which is defined for logistic regression as the deviance (Friedman 2001). Boosting iteratively upweights or ‘boosts’ observations that are misclassified at the previous iteration(s) and then applies a decision stump (a classification tree with one predictor or split) to the data, creating a succession of decision stumps

which are combined to create a final classifier. The number of decision stumps or ‘weak learners’ was set to be 6,500, which was estimated using tenfold cross-validation on our data. We selected SNPs based on the greatest relative influence of boosted estimates as defined by Friedman (2001):

$$\hat{J}_j^2 = \sum_{\text{split}_{x_j}} I_t^2$$

where I_t^2 is the improvement in reduction of the deviance observed from the split using the predictor x_j .

Functional magnetic resonance imaging methodologies

Subjects

Two hundred and sixty (CIT rs3847960 x CIT rs440299), 217 (DISC1 rs1411771 x CIT rs10744743) and 237 (NDEL1 rs4791707 x CIT rs10744743) healthy volunteers from the CBDB/NIMH Study of Biological Mechanisms of Genetic Association with Schizophrenia were used to explicitly test the epistatic effects of the (a) CIT × CIT, (b) DISC1 × CIT, and NDEL1 × CIT interactions outlined by the case–control study of schizophrenia risk on pre-frontal information processing efficiency during the *N*-back working memory task using the BOLD physiological response (for details on study subjects see Supplementary Table 5). Within these imaging samples, all SNPs were in Hardy–Weinberg equilibrium ($p < 0.05$) and no subgroup differed by age, IQ, and importantly, by 2-back accuracy or reaction time. Not confounded by individual performance, the relative amount of PFC activation is a measure of how individuals or genotype groups efficiently handle information—namely, inefficient subjects or groups (like patients with schizophrenia) will engage brain resources as reflected in higher activation without accompanying improvements in performance.

Task and functional image processing

Participants performed the *N*-back task (Callicott et al. 1999, 2003a) block fMRI paradigm that alternated between a working memory condition, 2-back, and a 0-back control condition. Whole brain BOLD fMRI data were acquired on a GE Signa 3T scanner (GE Systems; Milwaukee, WI) with a GE-EPI pulse sequence acquisition (24 contiguous axial slices of dimensions $3.75 \times 3.75 \times 6$ mm; flip angle 90° ; TR/TE 2,000/30 ms; FOV-24 cm; matrix 64×64 voxels). Images were processed with SPM5 software (<http://www.fil.ion.ucl.ac.uk/>) with realignment and correction for movement artifacts, spatial normalization in a standard stereotactic space (MNI template), smoothing with a 8 mm

full width half maximum Gaussian filter. First level images for each subject were created by modeling the two experimental conditions (2B and 0B) as boxcars convolved with a canonical hemodynamic response. A contrast image for the 2B > 0B contrast was estimated for each subject. These contrast images were used for a second-level random effect analysis.

Genotype effects were measured using multiple regression including main effects for each SNP, an interaction term, and variables to control for the effects of age and gender. As in prior reports utilizing the *N*-back, efficiency effects due to genotypic variation are sought a priori within bilateral DLPFC. Results are thresholded initially at $p < 0.001$, uncorrected, but we only report those foci surviving statistical control for false positives based on our prior hypotheses—in other words, a small volume correction within DLPFC at $p < 0.05$ FWE (Meyer-Lindenberg et al. 2008). This statistical correction approach to genetic association with brain physiology has been shown to be highly robust to false-positive results (Meyer-Lindenberg et al. 2008). To affirm that maximum deleterious effects on PFC efficiency occurred in individuals carrying risk alleles at both SNPs interacting clinically, values representing relative fMRI signal were extracted from significant loci and analyzed and displayed in SPSS.

Results

Single gene association results

Of the 12 SNPs genotyped in DISC1, rs1538976 minor allele homozygotes showed a trend for nominal association with schizophrenia (OR = 0.14; 95% confidence interval (CI) (0.02, 1.25), p value = 0.08) (Supplementary Table 6). One SNP in CIT was significantly associated with schizophrenia (rs10744743, heterozygote OR = 0.45, 95% CI (0.28, 0.71), p value = 0.001; minor allele homozygote OR = 0.22, 95% CI (0.13, 0.37), p value < 0.001) (Supplementary Table 7). In PAF1B1, a single SNP was significantly associated with schizophrenia (rs12938775; heterozygote OR = 0.62, 95% CI (0.41, 0.93), p value = 0.021; minor allele homozygote OR = 0.59, 95% CI (0.36, 0.95), p value = 0.030) (Supplementary Table 8). Minor allele carriers of the NDEL1 SNP rs4791707 were significantly less likely to be cases (heterozygote OR = 0.60, 95% CI (0.42, 0.86), p value = 0.005; minor allele homozygote OR = 0.56, 95% CI (0.33, 0.94), p value = 0.009) (Supplementary Table 8). SNPs in NDEL1 and FEZ1 did not show significant association with schizophrenia (Supplementary Tables 8 and 9).

Epistasis results

Figure 1 shows a schematic describing the building of forests of decision trees; see “Methods” for further discussion. Consistency in rankings of variable influence was observed between RF, GBM, and MCLR methods (Table 1). Empirical p values were obtained by randomly permuting case status in 500 replicates and comparing the observed variable importance measures with the distribution of variable importance values from the null replicates, which is equivalent to an experiment-wise empirical p value. SNPs were recoded as two variables for use with MCLR: one as minor allele dominant and one as minor allele recessive (see “Methods”). Seven SNPs had empirical p values <0.005 as being important interactors across at least two MLAs: CIT rs10744743, rs440299, rs203340, rs3847960, DISC1 rs1411771, PAFAH1B1 rs7212450 and NDEL1 rs4791707. Both the dominant and recessive coding of rs10744743 in CIT was ranked within the top 10 most influential variables using RF and MCLR. The same pattern was observed for all three MLAs for rs440299 also within CIT, which was not significantly associated with schizophrenia in the single SNP analysis. These two SNPs are in moderate LD with one another ($D' = 0.73$, $r^2 = 0.29$). A third SNP in CIT, under both dominant and recessive coding, was selected by MCLR as influential (rs3847960), and this SNP is in strong LD with rs10744743 ($D' = 0.93$, $r^2 = 0.85$) and moderate LD with rs440299 ($D' = 0.69$, $r^2 = 0.26$). Both GBM and RF also ranked this SNP (under the dominant coding) in the top 10 most influential variables. One additional SNP, rs203340, in CIT was ranked influential by MCLR and RF. Both GBM and RF detected rs7212450 in PAFAH1B1 which was not significantly associated with schizophrenia status in the single SNP analysis. The SNP in NDEL1 (rs4791707) that

was associated with schizophrenia in single SNP analyses was also selected as being influential across all MLAs. One SNP in DISC1, rs1411771, was ranked as having high influence using all three MLAs; this SNP did not show evidence for association in single SNP analyses.

We subjected these seven SNPs to all-possible 2-SNP interaction modeling using likelihood ratio tests (LRTs) for nested unconditional logistic regression models. Of the 21 interaction models tested, 4 showed significant evidence for interaction at the $p \leq 0.05$ level and 1 of the 4 models passed Bonferroni correction for the 21 tests: NDEL1 rs4791707 major allele homozygotes-CIT rs10744743 major allele homozygotes were 4.44 times more likely to be cases (95% CI (2.22, 8.88), LRT p value = 0.00013) than those carrying minor alleles at either SNP. An interaction between major allele homozygotes at DISC1 rs1411771 and the same SNP and genotype in CIT was also observed (OR = 3.07 (1.37, 6.98), LRT p value = 0.007). Two interactions between SNPs in CIT also showed evidence for interaction: rs3847960 minor allele carriers-rs440299 major allele homozygotes (OR = 2.16 (1.04, 4.46), LRT p value = 0.038) and rs3847960 minor allele carriers-rs203332 major allele homozygotes (OR = 2.90 (1.45, 5.79) LRT p value = 0.0030).

Statistical tests of epistasis in independent GWAS data

We sought to replicate our epistasis findings in two independent GWAS datasets: one consisting of a sample from Aberdeen and a sample from Germany (n cases = 1,221, n controls = 1,206) as described by Need et al. (2009) and the GAIN (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000021.v2.p1) study (n cases = 1,172, n controls = 1,378). Replication in the Need et al.

Table 1 Influence rankings and empirical p values of SNPs in DISC1 and protein interaction partners

Machine learning algorithm								
MCLR			GBM			RF		
Gene	SNP: coding	Empirical p value	Gene	SNP: coding	Empirical p value	Gene	SNP: coding	Empirical p value
CIT	rs10744743: D	<0.002	CIT	rs440299: D	<0.002	CIT	rs10744743: D	<0.002
CIT	rs440299: D	<0.002	NDEL1	rs4791707: D	<0.002	CIT	rs3847960: D	<0.002
CIT	rs3847960: R	<0.002	CIT	rs440299: R	<0.002	CIT	rs440299: D	<0.002
DISC1	rs1411771: D	<0.002	DISC1	rs1411771: D	<0.002	CIT	rs10744743: R	<0.002
NDEL1	rs4791707: D	<0.002	PAFAH1B1	rs7212450: D	0.12	NDEL1	rs4791707: D	<0.002
CIT	rs203340: D	<0.002	CIT	rs203332: R	0.11	CIT	rs203340: D	<0.002
CIT	rs203332: D	<0.002	DISC1	rs999710: R	0.27	CIT	rs440299: R	<0.002
CIT	rs10744743: R	<0.002	CIT	rs278109: D	0.35	DISC1	rs1411771: D	<0.002
CIT	rs3847960: D	<0.002	CIT	rs203340: R	0.29	PAFAH1B1	rs7212450: D	<0.002

D dominant coding of the minor allele, *R* recessive coding of the minor allele

(2009) study was performed using DISC1 rs1411771 and SNPs that were in strong LD with our other SNPs participating in epistasis where the exact same SNPs were not typed as part of the GWAS. In the German sample, using the same DISC1 SNP and rs4767848 as a proxy for CIT rs10744743 (HapMap CEU $D' = 1.0$, $r^2 = 1.0$ between the two SNPs), we observed a non-significant trend for evidence for epistasis (LRT p value = 0.051; OR = 1.35; 95% CI 0.95, 1.94), although the risk genotype for DISC1 was minor allele carriers instead of major allele homozygotes. In the pooled German and Aberdeen datasets described by Need et al. (2009) and in the GAIN dataset we did not observe evidence for epistasis at these SNPs.

Neuroimaging results

We tested for interactions within imaging space for DISC1 rs1411771 \times CIT rs10744743, CIT rs3847960 \times CIT rs203332, CIT rs3847960 \times CIT rs440299 and NDEL1 rs4791707 \times CIT rs10744743 using BOLD fMRI data from healthy subjects. Three of the four genetic interactions (rs1411771–rs10744743, rs3847960–rs440299, rs4791707–rs10744743) produced a significant interaction in imaging wherein subjects carrying clinical risk-associated alleles for both loci were most inefficient within dorsal prefrontal cortex—in essence (and as described for schizophrenic subjects and their unaffected siblings (Callicott et al. 2003a, b)) greater dlPFC activation with no concomitant advantage in either accuracy or reaction time, implicating a relative diminution in the neural tuning of cortical circuitry engaged during this task. We did not observe any statistically significant interaction between CIT SNPs rs3847960 and rs203332 (data not shown).

Within CIT, rs3847960 \times rs440299 produced maximal inefficiency within right Brodmann area (BA) 10 at (Talairach coordinate $x = 30$, $y = 37$, $z = 8$), surviving $p < 0.05$ small volume correction (SVC) in imaging space and with extracted signal values showing epistasis in statistical space ($F(1,125) = 9.96$, $p = 0.002$; Fig. 2). Between CIT and DISC1 (rs10744743 \times rs1411771), we found maximal inefficiency within the right middle frontal gyrus (BA9) again surviving $p < 0.05$ SVC in imaging space and significant in statistical space ($F(1,213) = 5.3$, $p < 0.05$; Fig. 3). Finally, for the most significant clinical interaction identified by machine learning, we found that NDEL1 (rs4791707) \times CIT (rs10744743) produced a maximal inefficiency effect within left middle frontal gyrus (BA 46) ($F(1,233) = 4.4$, $p < 0.05$, corrected; Fig. 4). As a post hoc exploration, we tested for interaction in the imaging data between the SNPs identified in the MLAs for DISC1 and NDEL1 against the single SNP showing association in PAFAH1B1, given the well-described molecular complex formed by these proteins in vivo. We found no

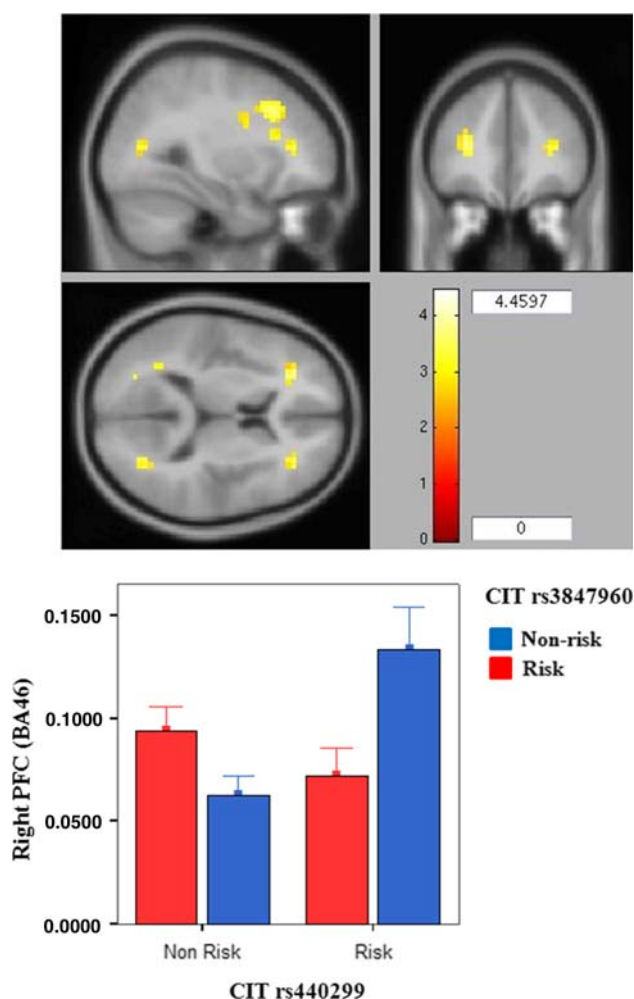


Fig. 2 CIT intragenic neuroimaging epistasis. CIT rs3847960 by CIT rs440299 interaction in normal subjects studied with BOLD fMRI during the N back working memory task. Figure at the *top* shows loci (in yellow) within brain in which significant interaction is found ($p < 0.05$ corrected) for all voxels within prefrontal cortical region of interest. Image at *bottom* shows relative degree of activation of DLPFC region showing imaging interaction based on genotypes at these two SNPs [fMRI signal extracted from maximum voxel and run as ANOVA in SPSS yielded $F(1,256) = 9.960$, $p = 0.002$]. The combination of both risk associated genotypes is disproportionately inefficient, i.e., have greatest activation without any difference in performance. Error bars represent 1 standard error of the mean. Frontal lobe (sub-gyral) mean activation extracted from 10 mm sphere at (30 37 8) Talairach

interaction for NDEL1 \times PAFAH1B1 (rs7212450), but identified a highly significant interaction between DISC1 (rs1411771) \times PAFAH1B1 (rs7212450) that produced an inefficiency signal within right middle frontal gyrus (BA 10) ($F(1,286) = 6.9$, $p = 0.009$, corrected).

Discussion

We report evidence that SNPs in three genes in the putative DISC1 pathway, DISC1, CIT and NDEL1, act in epistasis

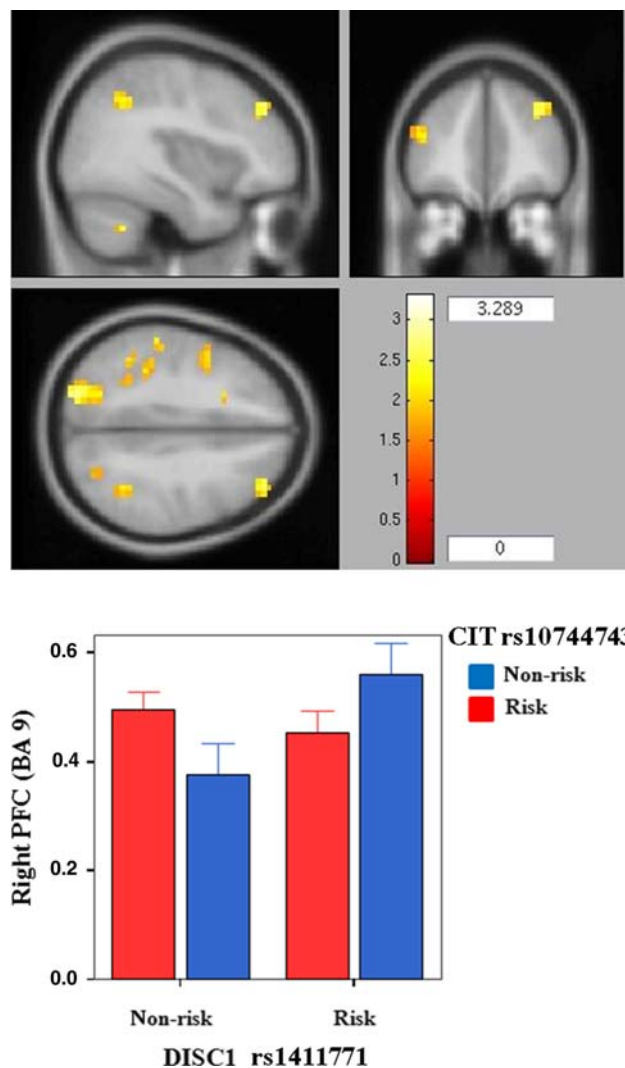


Fig. 3 DISC1 by CIT neuroimaging epistasis. DISC1 rs1411771 by CIT rs10744743 interaction in 217 normal subjects studied with BOLD fMRI during the *N*-back working memory task. The cross-sectional brain images at top figure show loci (in yellow) within brain showing a significant inefficiency effect associated with DISC1 risk \times CIT risk SNPs ($p < 0.05$ small volume correction SVC). The graph at bottom compares a measure of fMRI activation during this task for each genotype combination extracted from right PFC yielding a significant epistatic interaction [fMRI signal extracted from maximum voxel and run as ANOVA in SPSS yielded $F(1,213) = 5.3$ $p < 0.05$]

to influence risk for schizophrenia in our clinical sample, detected using MLAs. Further, we biologically validated three of the four significant interactions via neuroimaging in healthy controls; carriers of the combinations of schizophrenia risk-associated genotypes showed less efficient cognitive processing, similar to schizophrenia patients, than those carrying no risk-associated genotypes during a test of working memory. In addition, we observed non-significant evidence for interaction between DISC1 and CIT in an independent sample from Germany, although

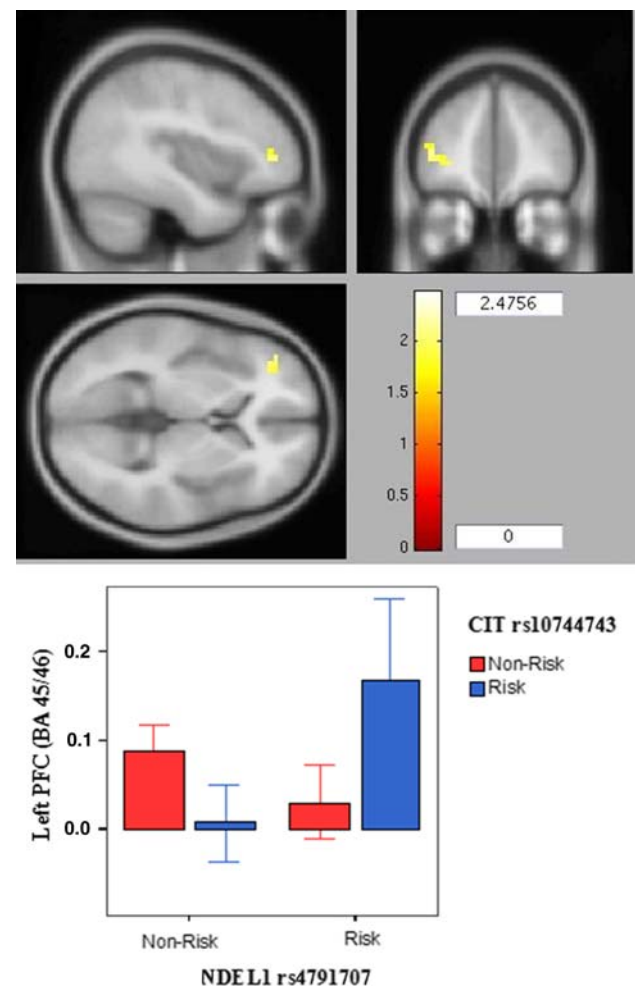


Fig. 4 NDEL1 by CIT imaging epistasis. NDEL1 rs4791707 by CIT rs10744743 interaction in 237 normal subjects studied with BOLD fMRI during the *N*-back working memory task. The cross-sectional brain images at top show loci (in yellow) within brain showing a significant inefficiency effect associated with NDEL1 risk \times CIT risk SNPs ($p < 0.05$ small volume correction SVC). The graph at bottom compares a measure of fMRI activation during this task for each genotype combination extracted from left PFC yielding a significant epistatic interaction [fMRI signal extracted from maximum voxel and run as ANOVA in SPSS yielded $F(1,233) = 4.4$ $p < 0.05$]

we were not able to directly replicate our interaction findings in the GAIN sample.

DISC1 is thought to be a scaffold protein hub that interacts physically with several partners (Porteous et al. 2006). Evidence for protein–protein interaction between DISC1 and NDEL1 has been reported using yeast two-hybrid screening, co-transformation, yeast mating and biosensor assays, with the translocation within DISC1 interrupting the ability of the NDEL1 protein to bind to DISC1 (Morris et al. 2003; Ozeki et al. 2003); the NDEL1 SNP we observed acting in epistasis is near (2,795 bp) the region encoding the NDEL1 domain that interacts with DISC1. Recent work has shown that neurite outgrowth

involves an interaction between *DISC1* and *NDEL1* proteins (Kamiya et al. 2006), and that statistical interaction in risk for schizophrenia has been reported between *DISC1*/*NDEL1* and *DISC1*/*NDE1* (Burdick et al. 2008) and also within *DISC1* (Schumacher et al. 2009). Since we did not genotype the same SNPs in *NDE1* and *NDEL1* reported to statistically interact with the functional *DISC1* SNP rs821616, we could not directly attempt replication of Burdick et al. (2008). However, we did not observe low-level (2-SNP) interaction between *NDE1* or *NDEL1* and *DISC1*, although we cannot exclude the possibility of high-order interactions. The *DISC1* SNP selected by MLAs (rs1411771) is not in LD ($D' = 0.28$, $r^2 = 0.01$, respectively) with the functional SNP rs821616 that has been associated with schizophrenia (Callicott et al. 2005) and bipolar disorder (Maeda et al. 2006), and thus could be in LD with an independent risk variant within *DISC1*; further, this 3' UTR SNP may have functional properties as it is predicted to be an exonic splicing enhancer (PupaSNP: <http://pupasnp.bioinfo.cipf.es/>) (Conde et al. 2006). In addition, rs1411771 has been reported to be part of a haplotype (with rs821616) associated with bipolar disorder (Palo et al. 2007), although other studies did not observe association with schizophrenia (Hennah et al. 2003) or with bipolar disorder/schizophrenia (Thomson et al. 2005), and another SNP in the 3' UTR (rs3737597) has been associated with schizophrenia in three independent Scandinavian case-control studies (Saetre et al. 2008). CIT has been suggested as a *DISC1* protein-interaction partner using yeast two-hybrid screening of human cDNA libraries (Morris et al. 2003), and is the human homolog of the *C. elegans* Citron, involved in the Rho signaling pathway which is important in axonal outgrowth (Bloom and Horvitz 1997; Furuyashiki et al. 1999; Morris et al. 2003). CIT knockout mice show reductions in hippocampal micro-neurons and developmental dysregulation in the central nervous system (Di Cunto et al. 2000). Interestingly, the SNPs selected by MLAs in CIT are near two SNPs (approximately 4,000 to 8,300 bp away) associated with bipolar disorder (Lyons-Warren et al. 2005) and are near exons in CIT that encode regions in CIT that interact with the *DISC1* protein. These CIT SNPs associated with and/or acting in epistasis to influence risk for schizophrenia and cognitive processing along with the two SNPs associated with bipolar disorder (Lyons-Warren et al. 2005) reside in a de novo gain-of-copy number region found in sporadic schizophrenia cases (Xu et al. 2008).

We also report significant association with two SNPs in *PAFAH1B1* but observed no evidence for two-way interaction between *PAFAH1B1* and other SNPs tested in the clinical data, although we did observe significant interaction between *PAFAH1B1* and *DISC1* using neuroimaging. Mutations in the gene *PAFAH1B1* result in lissencephaly,

which is characterized by deficits in neuronal migration (Fogli et al. 1999). The protein products of *DISC1*, *NDEL1* and *PAFAH1B1* form a trimolecular complex (Brandon et al. 2004); *PAFAH1B1* mRNA has been shown to be reduced in the hippocampus of schizophrenic versus control individuals; and schizophrenia-associated SNPs in *DISC1* are associated with *PAFAH1B1* mRNA expression in the hippocampus of individuals with schizophrenia (Lipska et al. 2006), suggesting variation in *PAFAH1B1* and/or interaction between *PAFAH1B1* and *DISC1* may influence risk for schizophrenia.

Our sample size was limited to detect interaction using traditional methods. We estimated power for two-SNP interactions in an unmatched case-control study using the GGIPOWER package (SJ3-1, st0032) for STATA 8.2 (College Park, TX) (Brown et al. 1999; Longmate 2001; Self et al. 1992). Assuming a minor risk allele frequency of 0.50, Fig. 5 shows estimated power under three types of two-SNP interactions using 300 cases and 300 controls, disease prevalence of 1%, alpha set at 0.05, and main effects for both loci set at either 1.0 (no main effect) or at 1.25 (weak main effect). The OR for the interaction varied from expected under the null hypothesis of no interaction ($\beta_{\text{int}} = 0$) by 1.0 (no interaction) to 10.0. Power to detect interaction was highest under recessive-recessive models; with an approximately 3.5-fold increased risk (compared to that expected under log additivity), we would have had 80% power to detect the interaction with our sample size. Under a dominant-recessive model, power was almost as high as under the recessive-recessive model. Under a dominant-dominant model we would expect 70% power if the interaction OR was approximately five times larger than that expected under log additivity where both single SNPs had no main effect (i.e., their individual ORs = 1.0). In the context of our observed interaction effect sizes, we would have had less than optimal power to detect interactions at the 80% or higher level using traditional logistic regression approaches.

In conclusion, our results suggest that interaction between polymorphisms in *DISC1*, CIT and *NDEL1*, which have been shown to interact biologically, influence risk for schizophrenia in our case-control sample and the same combinations of risk-associated genotypes are associated with less efficient cognitive processing in healthy controls. In addition, the SNPs detected using MLAs are either physically proximal to regions encoding *DISC1* protein interaction domains or, in the case of *DISC1*, predicted to be an exonic splicing enhancer, providing regions of interest for follow-up through additional genotyping or sequencing. This work adds to literature supportive of statistical interaction between candidate genes for schizophrenia or between candidate genes for schizophrenia and environmental risk factors (Burdick et al. 2008); in

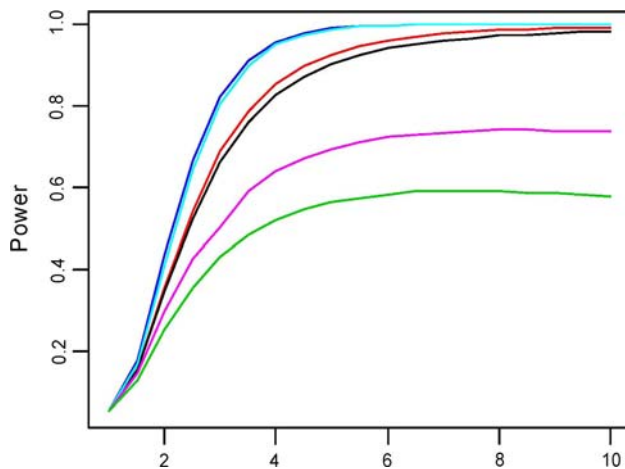


Fig. 5 Power for two-SNP interaction models. Models are as follows: *red* dominant–recessive model with main effect ORs = 1.0, *black* dominant–recessive model with main effect ORs = 1.25, *blue* recessive–recessive model with main effect ORs = 1.0, *turquoise* recessive–recessive model with main effect ORs = 1.25, *fuchsia* dominant–dominant model with main effect ORs = 1.0, *green* dominant–dominant model with main effect ORs = 1.25. Interaction is deviation from log additivity of the two genotypic coefficients

addition, the present study adds biologic validation to epistasis detected statistically and validated via an independent in vivo biological assay of brain function (Meyer-Lindenberg and Weinberger 2006).

Acknowledgments This study utilized the high performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD (<http://biowulf.nih.gov>). We thank Michael Dean, Bert Gold and Kate McGee for assistance with GAIN genotyping data. The authors declare the following competing interest: Pierandrea Muglia is a full-time employee of the pharmaceutical company GlaxoSmithKline who have filed patent applications for SNPs associated with schizophrenia (United States Patent Applications 20080176239 and 20080176240 and International Application Number PCT/EP2008/050477).

References

Blokland GA, McMahon KL, Hoffman J, Zhu G, Meredith M et al (2008) Quantifying the heritability of task-related brain activation and performance during the N-back working memory task: a twin fMRI study. *Biol Psychol* 79:70–79

Bloom L, Horvitz HR (1997) The *Caenorhabditis elegans* gene *unc-76* and its human homologs define a new gene family involved in axonal outgrowth and fasciculation. *Proc Natl Acad Sci USA* 97:3414–3419

Brandon NJ, Handford EJ, Schurov I, Rain JC, Pelling M et al (2004) Disrupted in schizophrenia 1 and Nudel form a neurodevelopmentally regulated protein complex: implications for schizophrenia and other major neurological disorders. *Mol Cell Neurosci* 25:42–55

Breiman L (2001) Random forests. *Mach Learn* 45:5–32

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth, Belmont

Brown BW, Lovato J, Russell K (1999) Asymptotic power calculations: description, examples, computer code. *Stat Med* 18:3137–3151

Burdick KE, Kamiya A, Hodgkinson CA, Lencz T, DeRosse P et al (2008) Elucidating the relationship between DISC1, NDEL1, and NDE1 and the risk for schizophrenia: evidence of epistasis and competitive binding. *Hum Mol Genet* 17:2462–2473

Callicott JH, Mattay VS, Bertolino A, Finn K, Coppola R et al (1999) Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cereb Cortex* 9:20–26

Callicott JH, Mattay VS, Verchinski BA, Marenco S, Egan MF et al (2003a) Complexity of prefrontal cortical dysfunction in schizophrenia: more than up or down. *Am J Psychiatry* 160:2209–2215

Callicott JH, Egan MF, Mattay VS, Bertolino A, Bone AD et al (2003b) Abnormal fMRI response of the dorsolateral prefrontal cortex in cognitively intact siblings of patients with schizophrenia. *Am J Psychiatry* 160:709–719

Callicott JH, Straub RE, Pezawas L, Egan MF, Mattay VS et al (2005) Variation in DISC1 affects hippocampal structure and function and increases risk for schizophrenia. *Proc Natl Acad Sci USA* 102:8627–8632

Cannon TD, Hennah W, van Erp TG, Thompson PM, Lonnqvist J et al (2005) Association of DISC1/TRAX haplotypes with schizophrenia, reduced prefrontal gray matter, and impaired short- and long-term memory. *Arch Gen Psychiatry* 62:1205–1213

Carmargo LM, Collura V, Rain JC, Mizuguchi K, Hermjakob H et al (2007) Disrupted in Schizophrenia 1 interactome: evidence for the close connectivity of risk genes and a potential synaptic basis for schizophrenia. *Mol Psychiatry* 12:74–86

Clapcote SJ, Lipina TV, Millar JK, Mackie S, Christie S et al (2007) Behavioral phenotypes of Disc1 missense mutations in mice. *Neuron* 54:387–402

Conde L, Vaquerizas JM, Dopazo H, Arbiza L, Reumers J et al (2006) PupaSuite: finding functional single nucleotide polymorphisms for large-scale genotyping purposes. *Nucleic Acids Res* 34:W621–W625

de Bakker PIW, Yelensky R, Pe'er I, Gabriel SB, Daly MJ et al (2005) Efficiency and power in genetic association studies. *Nat Genet* 37:1217–1223

Devon RS, Anderson S, Teague PW, Burgess P, Kipari TM et al (2001) Identification of polymorphisms within Disrupted in Schizophrenia 1 and Disrupted in Schizophrenia 2, and an investigation of their association with schizophrenia and bipolar affective disorder. *Psychiatr Genet* 11:71–78

Di Cunto F, Imarisio S, Hirsch E, Broccoli V, Bulfone A et al (2000) Defective neurogenesis in citron kinase knockout mice by altered cytokinesis and massive apoptosis. *Neuron* 28:115–127

Egan MF, Goldberg TE, Gscheidle T, Weirich M, Bigelow LB et al (2000) Relative risk of attention deficits in siblings of patients with schizophrenia. *Am J Psychiatry* 157:1309–1316

Feng Y, Olson EC, Stukenberg PT, Flanagan LA, Kirschner MW et al (2000) LIS1 regulates CNS lamination by interacting with mNude, a central component of the centrosome. *Neuron* 28:279–293

Fogli A, Guerrini R, Moro F, Fernandez-Alvarez E, Livet MO et al (1999) Intracellular levels of the LIS1 protein correlate with clinical and neuroradiological findings in patients with classical lissencephaly. *Ann Neurol* 45:154–161

Friedman JH (2001) Greedy function approximation: a gradient boosting machine. *Ann Stat* 29:1189–1232

Furuyashiki T, Fujisawa K, Fijita A, Madaule P, Uchino S et al (1999) Citron, a Rho-target, interacts with PSD-95/SAP-90 at glutamatergic synapses in the thalamus. *J Neurosci* 19:109–118

Hennah W, Varilo T, Kestilä M, Paunio T, Arajärvi R et al (2003) Haplotype transmission analysis provides evidence of

- association for DISC1 to schizophrenia and suggests sex-dependent effects. *Hum Mol Genet* 12:3151–3159
- Hennah W, Tomppo L, Hiekkalinna T, Palo OM, Kilpinen H et al (2007) Families with the risk allele of DISC1 reveal a link between schizophrenia and another component of the same molecular pathway, NDE1. *Hum Mol Genet* 16:453–462
- Hennah W, Thomson P, McQuillin A, Bass N, Loukola A et al (2008) DISC1 association, heterogeneity and interplay in schizophrenia and bipolar disorder. *Mol Psychiatry* 14:865–873 (Epub March 4)
- Hikida T, Jaaro-Peled H, Seshadri S, Oishi K, Hookway C et al (2007) Dominant-negative DISC1 transgenic mice display schizophrenia-associated phenotypes detected by measures translatable to humans. *Proc Natl Acad Sci USA* 104:14501–14506
- Hodgkinson CA, Goldman D, Jaeger J, Persaud S, Kane JM et al (2004) Disrupted in schizophrenia 1 (DISC1): association with schizophrenia, schizoaffective disorder, and bipolar disorder. *Am J Hum Genet* 75:862–872
- Hodgkinson CA, Goldman D, Ducci F, DeRosse P, Caycedo DA et al (2007) The FEZ1 gene shows no association to schizophrenia in Caucasian or African American populations. *Neuropsychopharmacology* 32:190–196
- Kähler AK, Djurovic S, Kulle B, Jönsson EG, Agartz I et al (2008) Association analysis of schizophrenia on 18 genes involved in neuronal migration: MDGA1 as a new susceptibility gene. *Am J Med Genet Part B (Neuropsychiatr Genet)* 147B:1089–1100
- Kamiya A, Tomoda T, Chang J, Takaki M, Zhan C et al (2006) DISC1-NDEL1/NUDEL protein interaction, an essential component for neurite outgrowth, is modulated by genetic variations of DISC1. *Hum Mol Genet* 15:3313–3323
- Kockelkorn TT, Arai M, Matsumoto H, Fukuda N, Yamada K et al (2004) Association study of polymorphisms in the 5' upstream region of human DISC1 gene with schizophrenia. *Neurosci Lett* 368:41–45
- Koga M, Ishiguro H, Horiuchi Y, Albalushi T, Inada T et al (2007) Failure to confirm the association between the FEZ1 gene and schizophrenia in a Japanese population. *Neurosci Lett* 417:326–329
- Kooperberg C, Ruczinski I (2005) Identifying interacting SNPs using Monte Carlo logic regression. *Genet Epidemiol* 28:157–170
- Kooperberg C, Ruczinski I, LeBlanc ML, Hsu L (2001) Sequence analysis using logic regression. *Genet Epidemiol* 21(Suppl 1):S626–S631
- Li W, Zhou Y, Jentsch JD, Brown RA, Tian X et al (2007) Specific developmental disruption of disrupted-in-schizophrenia-1 function results in schizophrenia-related phenotypes in mice. *Proc Natl Acad Sci USA* 104:18280–18285
- Lipska BK, Peters T, Hyde TM, Halim N, Horowitz C et al (2006) Expression of DISC1 binding partners is reduced in schizophrenia and associated with DISC1 SNPs. *Hum Mol Genet* 15:1245–1258
- Liu Y-L, Fann CS, Liu CM, Chen WJ, Wu JY et al (2006) A single nucleotide polymorphism fine mapping study of chromosome 1q42.1 reveals the vulnerability genes for schizophrenia, GNPAT and DISC1: Association with impairment of sustained attention. *Biol Psychiatry* 60:554–562
- Longmate JA (2001) Complexity and power in case-control association studies. *Am J Hum Genet* 68:1229–1237
- Lyons-Warren A, Chang JJ, Balkissoon R, Kamiya A, Garant M et al (2005) Evidence of association between bipolar disorder and Citron on chromosome 12q24. *Mol Psychiatry* 10:807–809
- Maeda K, Nwulia E, Chang J, Balkissoon R, Ishizuka K et al (2006) Differential expression of disrupted-in-schizophrenia (DISC1) in bipolar disorder. *Biol Psychiatry* 60:929–935
- Meyer-Lindenberg A, Weinberger DR (2006) Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nat Rev Neurosci* 7:818–827
- Meyer-Lindenberg A, Nicodemus KK, Egan MF, Calicott JH, Mattay V et al (2008) False positives in imaging genetics. *Neuroimage* 40:655–661
- Millar JK, Wilson-Annan JC, Anderson S, Christie S, Taylor MS et al (2000) Disruption of two novel genes by a translocation cosegregating with schizophrenia. *Hum Mol Genet* 9:1415–1423
- Millar JK, Christie S, Anderson S, Lawson D, Hsiao-Wei Loh D et al (2001) Genomic structure and localisation within a linkage hotspot of Disrupted in Schizophrenia 1, a gene disrupted by a translocation segregating with schizophrenia. *Mol Psychiatry* 6:173–178
- Millar JK, James R, Christie S, Porteous DJ (2005) Disrupted in schizophrenia 1 (DISC1): subcellular targeting and induction of ring mitochondria. *Mol Cell Neurosci* 30:477–484
- Miyoshi K, Honda A, Baba K, Taniguchi M, Oono K et al (2003) Disrupted-In-Schizophrenia 1, a candidate gene for schizophrenia, participates in neurite outgrowth. *Mol Psychiatry* 8:685–694
- Morris JA, Kandpal G, Ma L, Austin CP (2003) DISC1 (Disrupted-In-Schizophrenia 1) is a centrosome-associated protein that interacts with MAP1A, MIPT3, ATF4/5 and NUDEL: regulation and loss of interaction with mutation. *Hum Mol Genet* 12:1591–1608
- Need AC, Ge D, Weale ME, Maia J, Feng S et al (2009) A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet* 5:e1000373
- Nicodemus KK, Malley JD (2009) Predictor correlation impacts machine learning algorithms: implications for genomic studies. *Bioinformatics* 25:1884–1890
- Nicodemus KK, Wang W, Shugart YY (2007) Stability of variable importance scores and rankings using statistical learning tools on single-nucleotide polymorphisms and risk factors involved in gene \times gene and gene \times environment interactions. *BMC Proc* 1(Suppl 1):S58
- Numata S, Ueno S, Iga J, Nakataki M, Ohmori T et al (2008) No association between the NDE1 gene and schizophrenia in the Japanese population. *Schizophr Res* 99:367–369
- Ozeki Y, Tomoda T, Kleiderlein J, Kamiya A, Bord L et al (2003) Disrupted-in-Schizophrenia 1 (DISC-1): mutant truncation prevents binding to NudE-like (NUDEL) and inhibits neurite outgrowth. *Proc Natl Acad Sci USA* 100:289–294
- Palo OM, Anttila M, Silander K, Hennah W, Kilpinen H et al (2007) Association of distinct allelic haplotypes of DISC1 with psychotic and bipolar spectrum disorders and with underlying cognitive impairments. *Hum Mol Genet* 16:2517–2528
- Pletnikov MV, Ayhan Y, Nikolskaia O, Ovanesov MV, Huang H et al (2008) Inducible expression of mutant human DISC1 in mice is associated with brain and behavioral abnormalities reminiscent of schizophrenia. *Mol Psychiatry* 13:173–186
- Porteous DJ, Thomson P, Brandon NJ, Millar JK (2006) The genetics and biology of DISC1—an emerging role in psychosis and cognition. *Biol Psychiatry* 60:123–131
- Saetre P, Agartz I, De Franciscis A, Lundmark P, Djurovic S et al (2008) Association between a disrupted-in-schizophrenia (DISC1) single nucleotide polymorphism and schizophrenia in a combined Scandinavian case-control sample. *Schizophr Res* 106:237–241
- Sanders AR, Duan J, Levinson DF, Shi J, He D et al (2008) No significant association of 14 candidate genes with schizophrenia in a large European ancestry sample: implications for psychiatric genetics. *Am J Psychiatry* 165:497–506
- Schumacher J, Laje G, Jamra RA, Becker T, Mühleisen TW et al (2009) The DISC locus and schizophrenia—evidence from an association study in a central European sample and from a meta-analysis across different European populations. *Hum Mol Genet* 18:2719–2727 (Epub May 4)

- Self SG, Mauritsen RH, Ohara J (1992) Power calculations for likelihood ratio tests in generalized linear models. *Biometrics* 48:31–39
- Shu T, Ayala R, Nguyen MD, Xie Z, Gleeson JG et al (2004) Ndel operates in a common pathway with LIS1 and cytoplasmic dynein to regulate cortical neuronal positioning. *Neuron* 14:263–277
- Song W, Li W, Feng J, Heston LL, Scaringe WA et al (2008) Identification of high risk DISC1 structural variants with a 2% attributable risk for schizophrenia. *Biochem Biophys Res Commun* 367:700–706
- St. Clair D, Blackwood D, Muir W, Carothers A, Walker M et al (1990) Association within a family of a balanced autosomal translocation with major mental illness. *Lancet* 336:13–16
- Tarricone C, Perrina F, Monzani S, Massimiliano L, Kim MH et al (2004) Coupling PAF signaling to dynein regulation: structure of LIS1 in complex with PAF-acetylhydrolase. *Neuron* 44:809–821
- Taya S, Shinoda T, Tsuboi D, Asaki J, Nagai K et al (2007) DISC1 regulates the transport of the NUDEL/LIS1/14–3–3 ϵ complex through kinesin-1. *J Neurosci* 27:15–26
- Thomson PA, Wray NR, Millar JK, Evans KL, Hellard SL et al (2005) Association between the TRAX/DISC locus and both bipolar disorder and schizophrenia in the Scottish population. *Mol Psychiatry* 10:657–668
- Tomppo L, Hennah W, Lahermo P, Loukola A, Tuulio-Henriksson A et al (2009) Association between genes of disrupted in schizophrenia 1 (DISC1) interactors and schizophrenia supports the role of the DISC1 pathway in the etiology of major mental illnesses. *Biol Psychiatry* 65:1055–1062 (Epub February 27)
- Wood LS, Pickering EH, Dechairo BM (2007) Significant support for DAO as a schizophrenia susceptibility locus: examination of five genes putatively associated with schizophrenia. *Biol Psychiatry* 61:1195–1199
- Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA et al (2008) Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet* 40:880–885
- Yamada K, Nakamura K, Minabe Y, Iwayama-Shigeno Y, Takao H et al (2004) Association analysis of FEZ1 variants with schizophrenia in Japanese cohorts. *Biol Psychiatry* 56:683–690
- Zhang X, Tochigi M, Ohashi J, Maeda K, Kato T et al (2005) Association study of the DISC1/TRAX locus with schizophrenia in a Japanese population. *Schizophr Res* 79:175–180
- Zhang F, Sarginson J, Crombie C, Walker N, St Clair D et al (2006) Genetic association between schizophrenia and the DISC1 gene in the Scottish population. *Am J Med Genet Part B (Neuropsychiatr Genet)* 141B:155–159
- Zhang F, Sarginson J, Crombie C, Walker N, St Clair D et al (2007) Erratum: Genetic association between schizophrenia and the DISC1 gene in the Scottish population. *Am J Med Genet Part B (Neuropsychiatr Genet)* 144B:840